INF2178 Experimental Design For Data Science
Technical Assignment #4

Student Name: Liguangxuan He | Student ID: 1006141809

Background Information About This Data

The data set under review, a subset derived from a longitudinal study on MRI results among older individuals with and without dementia, forms part of the Open Access Series of Imaging Studies (OASIS). This initiative, a collaboration among renowned institutions and researchers, aims to propel neuroscience research by providing free access to brain MRI datasets. Encompassing 294 entries across 15 columns, the datasets primarily focuses on an elderly demographic, evident from the mean age of 76.41 years.

Data Cleaning

I completed three tasks for data preparation: initially, I removed the 'Unnamed: 0' column, then addressed missing values in the 'SES' and 'MMSE' columns by imputing their means, and finally, I transformed the 'Group', 'M/F', and 'Hand' columns into categorical data types to facilitate potential future analysis. Furthermore, I executed a loop to tally the frequencies within these categorical columns.

---

1. Exploratory Data Analysis (EDA)

Research Question 1: What are the strength and direction of relationships among quantitative features?

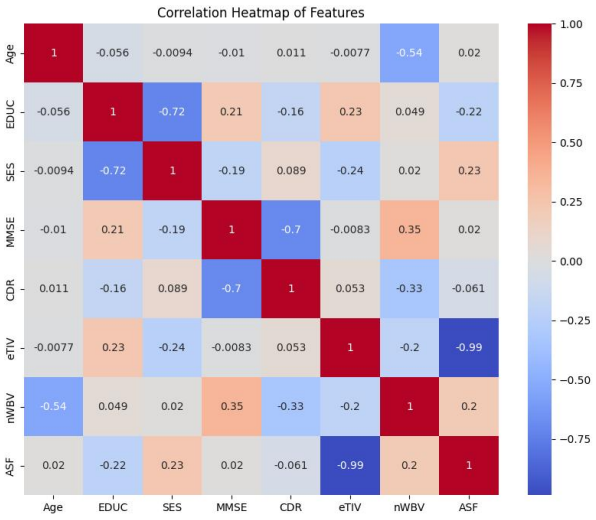


Figure 1. Correlation Heatmap of Features

Referring to Figure 1, the heatmap provides a visual representation of the correlation between different variables in the dataset. Notably, there is a strong negative correlation (r = -0.72) between education (EDUC) and socioeconomic status (SES), suggesting that higher levels of education are associated with lower SES in this sample. Conversely, there is a significant negative correlation (r = -0.7) between the Mini-Mental State Examination (MMSE) scores and the Clinical Dementia Rating (CDR), indicating that as cognitive impairment increases (higher CDR), cognitive function scores decrease (lower MMSE). Additionally, there is a strong negative correlation between normalized whole brain volume (nWBV) and age (r = -0.54), implying that brain volume decreases with age. The atlas scaling factor (ASF) is strongly negatively correlated with the estimated total intracranial volume (eTIV) (r = -0.99), which is expected as ASF is a scaling factor inversely related to eTIV. These correlations are crucial as they suggest potential relationships between demographic, socioeconomic, and clinical variables with brain structure and cognitive health in the elderly population under study.

Research Question 2: How do factors such as education, socioeconomic status, MMSE scores, CDR scores, brain volume, eTIV, and ASF differ among varying levels of dementia?

Refering to Figure 2, the use of boxplots depict how various metrics distribute across different dementia statuses—converted, demented, and nondemented. The estimated total intracranial volume (eTIV) shows slight variation across groups, while the normalized whole brain volume (nWBV) decreases notably in the demented group, indicating a possible relationship between brain atrophy and dementia. Education levels appear higher in the nondemented group, suggesting a

potential protective role of education against dementia. Socioeconomic status (SES) displays some overlap, yet there's a discernible decrease in the demented group. Mini-Mental State Examination (MMSE) scores decrease substantially from nondemented to demented, reinforcing its use as a cognitive impairment indicator. Clinical Dementia Rating (CDR) scores are higher in the demented group, which is consistent with its purpose to reflect severity of dementia. The Atlas Scale Factor (ASF) varies inversely with eTIV, with little variation between dementia statuses, which aligns with its mathematical relationship to eTIV.

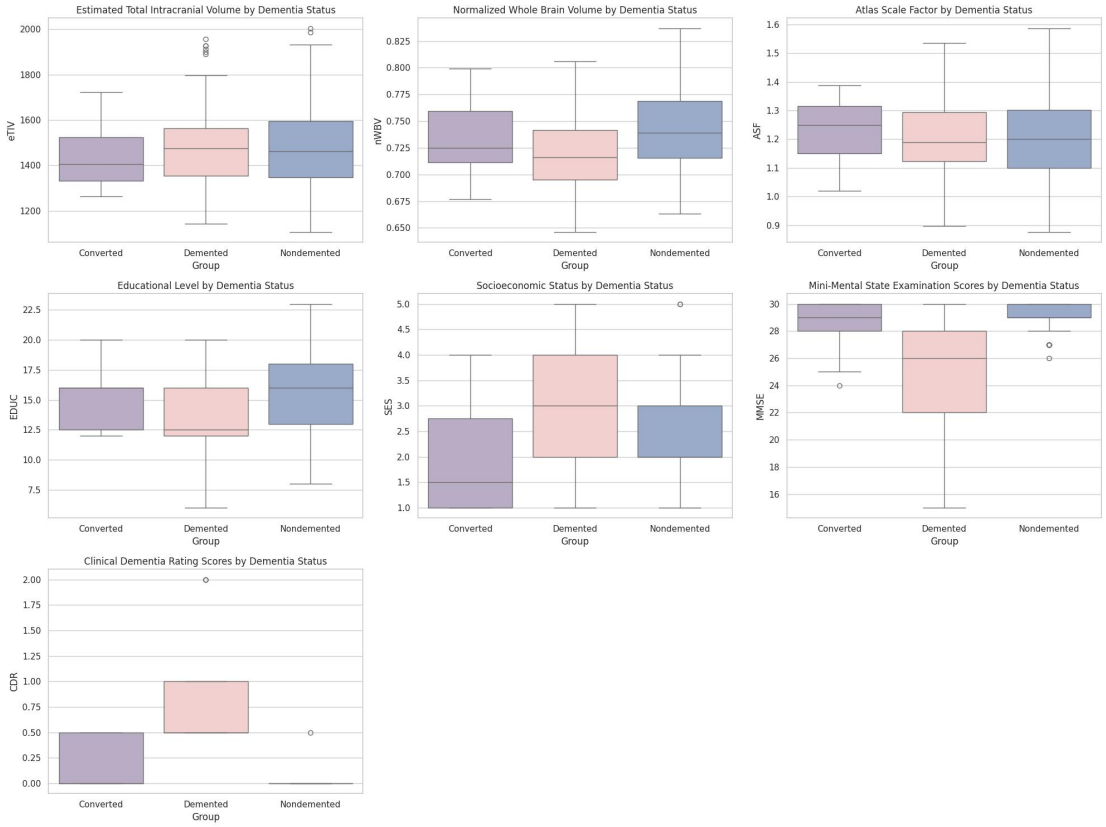


Figure 2. Boxplots for factors differ among dementia statuses

Research Question 3: How does age vary across different dementia statuses, and does gender influence age distribution within these groups?

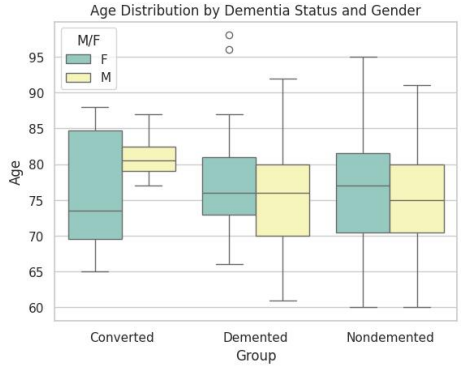


Figure 3. boxplots on age distribution by dementia status and gender

Referring to Figure 3, the boxplot reveals age distribution differences by gender within each dementia status group. Males in the 'converted' category generally skew younger compared to females. Across all groups, there's a broad age range, but 'demented' and 'nondemented' categories show more gender parity in age. Notably, outliers are present, indicating that there are individuals who fall well outside the typical age range, particularly in the 'converted' and 'demented' groups.

---

2. Mixed-effects ANOVA analysis (Within-and Between subjects)

Research Question: How do Mini-Mental State Examination (MMSE) scores change over multiple visits within individuals, and does this change differ by dementia status (nondemented, demented, converted)?

Table 1. ANOVA summary

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Group	1322.02	2	141	661.01	56.10	0.000	0.443	nan

Visit	21.53	1	141	21.53	8.53	0.004	0.057	1.000
Interaction	16.20	2	141	8.10	3.21	0.043	0.044	nan

Table 2. post hoc test
(Nondemented = Non, Demented= Dem, Converted=Con, True=T, False=F, Two-sided = two)

Constrast	Visit	A	B	Paired	Parametric	T	dof	alternative	p-unc
Visit	-	1	2	T	T	2.88	143.00	two	0.005
Group	-	Non	Dem	F	T	9.51	65.51	two	0.000
Group	-	Non	Con	F	T	1.30	12.32	two	0.216
Group	-	Dem	Con	F	T	-6.75	50.48	two	0.000
Visit * Group	1	Non	Dem	F	T	9.12	68.19	two	0.000
Visit * Group	1	Non	Con	F	T	-0.49	14.00	two	0.633
Visit * Group	1	Dem	Con	F	T	-8.08	60.17	two	0.000
Visit * Group	2	Non	Dem	F	T	8.48	66.15	two	0.000
Visit * Group	2	Non	Con	F	T	1.82	11.80	two	0.095
Visit * Group	2	Dem	Con	F	T	-4.52	33.37	two	0.000

Table 1, summarizing the mixed-effects ANOVA analysis, reveals that MMSE scores indeed vary significantly by dementia status ($p < 0.001$) with a substantial effect size ($np2 = 0.443$), suggesting a strong association between dementia status and cognitive performance across participants. Additionally, there is a significant within-subjects effect of visit ($p = 0.004$, $np2 = 0.057$), indicating that MMSE scores change over multiple visits. The interaction effect between group and visit is also significant ($p = 0.043$, $np2 = 0.044$), suggesting that the change in MMSE scores over visits differs depending on the dementia status of the participants.

Table 2 details the post hoc test results and provides further insights. The significant change in MMSE scores between the first and second visits ($p = 0.005$) suggests cognitive scores evolved over time within individuals. When comparing dementia statuses, the nondemented and demented groups show a significant difference in MMSE scores ($p < 0.001$), as do the demented and converted groups ($p < 0.001$). However, the difference between the nondemented and converted groups was not significant ($p = 0.216$). The interaction effects reflect these trends, with significant interactions between visit and group for nondemented versus demented in both visits ($p < 0.001$), but not for nondemented versus converted, indicating the pattern of cognitive change across visits varies by dementia status. These results highlight that dementia status significantly impacts cognitive trajectory over time, with distinct patterns observed between groups.

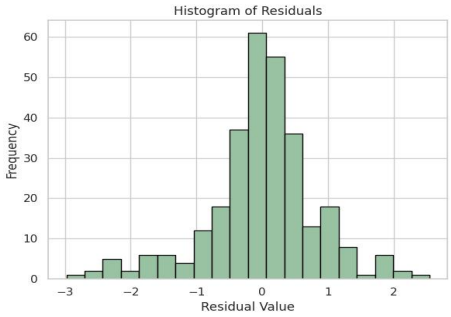
- Assumption Check - Normality of Residuals

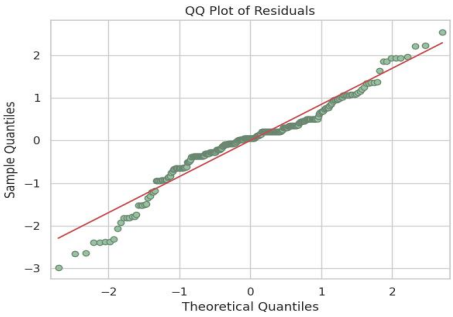
Table 3 indicates that the assumption of normality is violated for the residuals in most groups as the p-values are below the typical alpha level of 0.05, with the Shapiro-Wilk test statistic values ranging from 0.73 to 0.95. Specifically, 'Nondemented' visits 1 and 2, 'Demented' visit 2, and 'Converted' visit 1 show a significant deviation from normality ($p < 0.05$). Conversely, 'Converted' at visit 2 exhibits a p-value of 0.08, suggesting that the normality assumption might not be violated in this group-visit combination. Referring to Figure 4, while the histogram of residuals appears reasonably symmetric, the QQ plot reveals deviations from normality, particularly in the tails. This combination of tests and visual inspections suggests that while the ANOVA assumptions may not be fully met, particularly in the tails where outliers might exist, the central tendency of the residuals does not deviate drastically from normality.

Table 3. Shapiro-Wilk Test

Group	VIsit	T-Statistic	p-Value
Nondemented	1	0.80	0.00
Nondemented	2	0.81	0.00
Demented	1	0.95	0.01
Demented	2	0.90	0.00
Converted	1	0.73	0.00
Converted	2	0.88	0.08

Figure 4.Histogram and QQ plot for residuals





- Assumption Check - Homogeneity of Variances

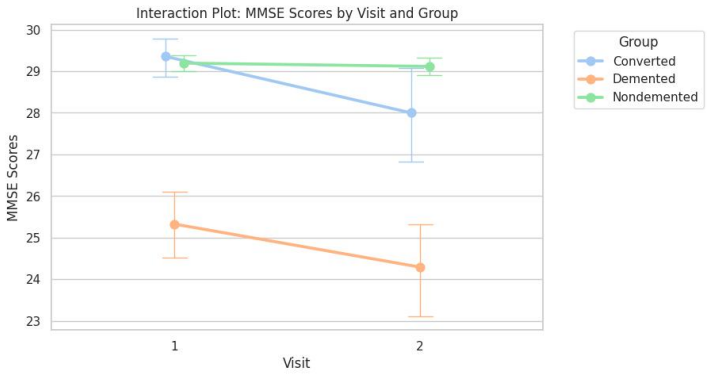
Table 4. Levene's Test

Parameter	Value
Test Statistic (W)	41.52
p value	0.00

The results from Table 4's Levene's Test, with a high test statistic of 41.52 and a p-value close to zero, suggest a violation of the homogeneity of variances assumption in the ANOVA analysis. A p-value significantly lower than the conventional alpha level of 0.05 indicates that the groups do not have equal variances, which is a critical condition for the validity of ANOVA results. This implies that the variability of scores within each group is not consistent across the different levels of the independent variable, which in this case are the dementia statuses.

- Interaction Plot

Figure 5. Interaction Plot

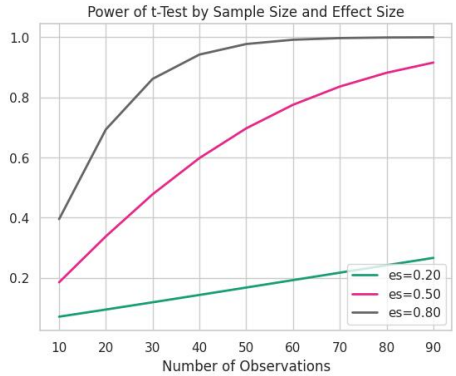


The interaction plot aligns with the previously discussed ANOVA findings by illustrating changes in MMSE scores across visits for each dementia status group. It shows a consistent trend in the nondemented group across visits, whereas the demented group exhibits a notable decline between the first and second visit. The converted group's scores also decrease, albeit less dramatically than the demented group. This visualization underscores the ANOVA's interaction effect, with the most substantial cognitive decline occurring in the demented group. The variability within groups, indicated by the plot's error bars, confirms the significant differences in cognitive score changes over time, as indicated by the visit on the x-axis.

---

3. Statistical Power Analysis

Figure 6. Power Analysis Graph



According to Figure 6, achieving a statistical power of 0.8 requires around 65 observations for a medium effect size (es=0.50) and approximately 25 observations for a large effect size (es=0.80). This suggests that the number of observations needed to achieve an acceptable level of power increases as the effect size decreases. With the ANOVA model's reported power of 0.934, it's clear that the study is well-powered, implying a high likelihood of detecting an effect, assuming one exists. The high power indicates a strong model that is not prone to a Type II error, affirming the reliability of the ANOVA results given the sample size and effect size reflected in the study.